



Liz Sexton-Kennedy

# CDF Computing



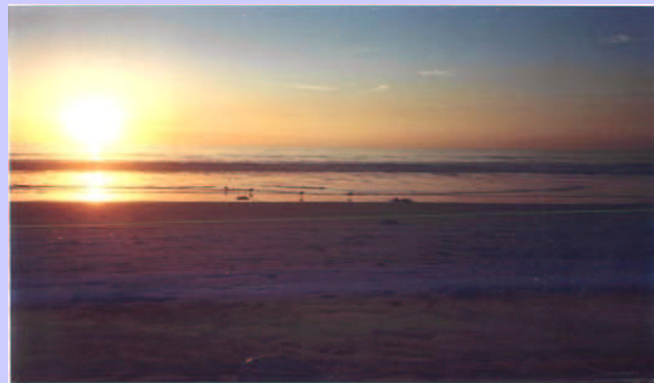
Toronto



INFN Bologna



Glasgow



UC San Diego



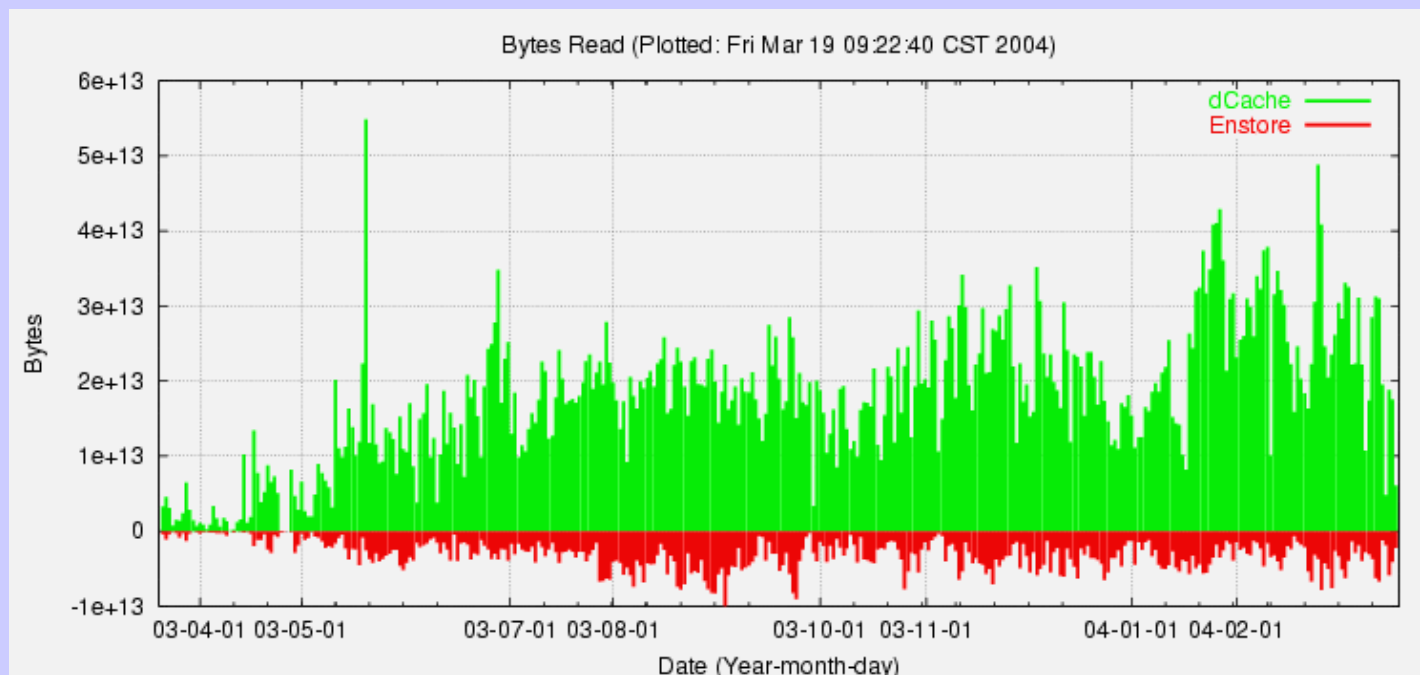
- ◆ Current Scale
- ◆ Plans for future
- ◆ What we've done
- ◆ What we want to do

All MC done at remote sites, and 2 remote CAFs



## ***CDF's Current Offline Scale***

- ◆ 685M unique events have been processed
- ◆ 747 TB on tape, 4 copies x 250kB/event + MC
- ◆ 48TB/day moved, at 900MB/sec sustained
- ◆ CDF has pioneered commodity file servers
- ◆ We have deployed 1382 CPUs and 318 TB of disk in our PC farms, CDF CAF is a big success!





## *Plans for the Future*

- ◆ Strong support from IFC & Bird reviews
- ◆ PAC agrees that plans for expanded DAQ bandwidth is well motivated by the physics
- ◆ Held a global computing workshop in Florida. A big success! 2<sup>nd</sup> round of CAF's by May 1st
- ◆ We are aiming for ~7 sites with ~200 CPUs each ~1400 CPUs at FNAL for users plus ~500 CPUs for reconstruction

*SAM is the enabling technology*

[http://cdfkits.fnal.gov/DIST/doc/DCAF/web/Florida\\_workshop.html](http://cdfkits.fnal.gov/DIST/doc/DCAF/web/Florida_workshop.html)

Karlsruhe



Barcelona



Japan



Taiwan



Korea



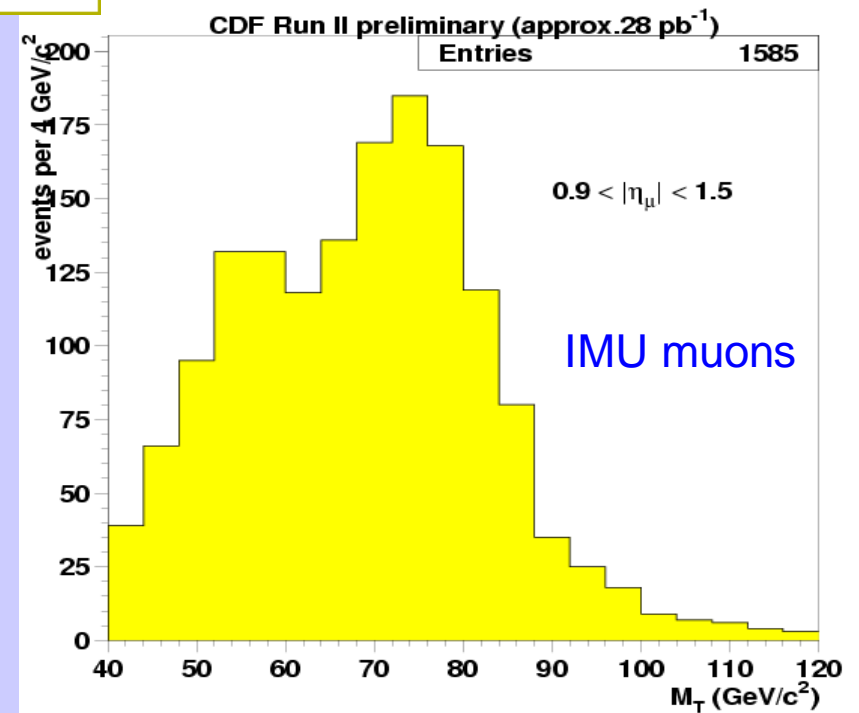
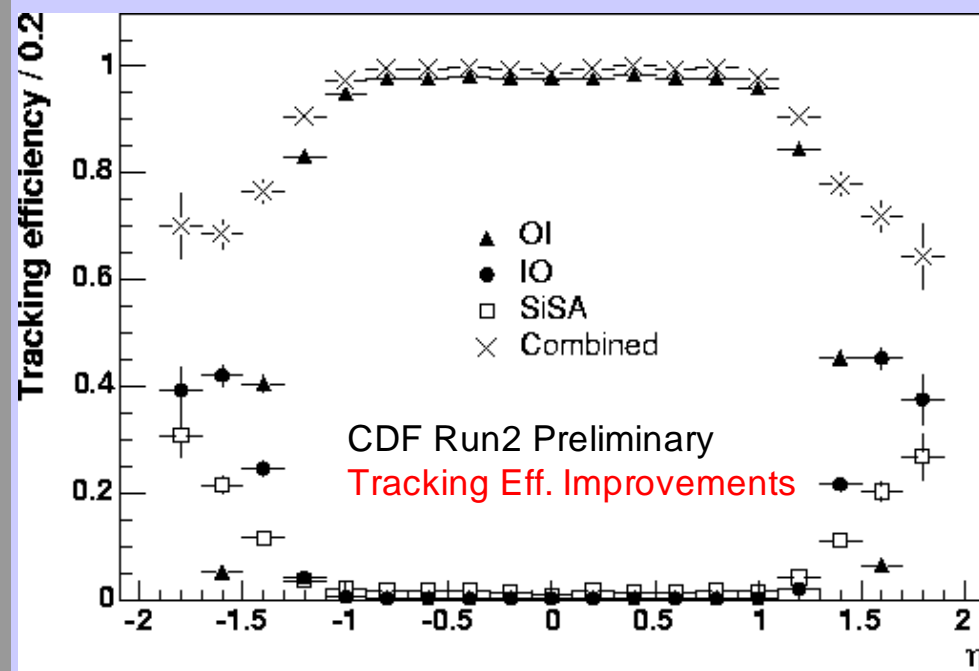
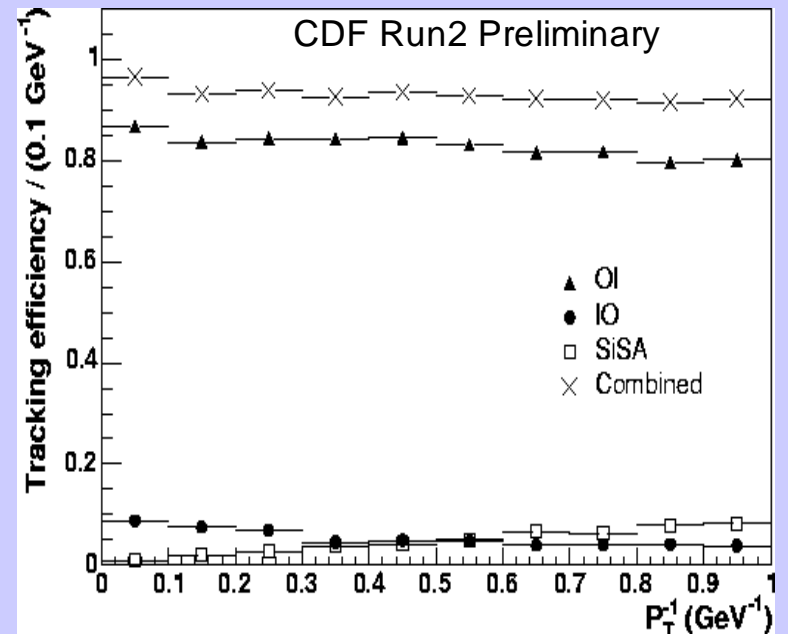
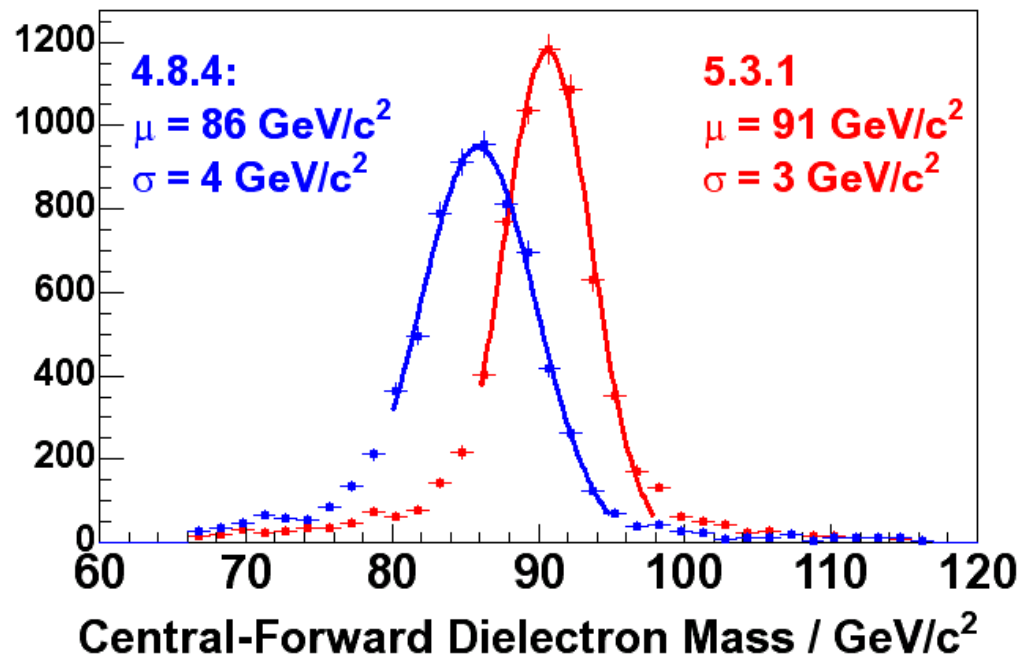


## ***What We've Done***

- ◆ Reconstruction passes(all data, 3 times):
  - ◆ 4.8.4 ran non-optimized, no forward tracking or IMU
  - ◆ 5.1.0 optimized, alignment improved, beamline used
  - ◆ 5.3.0 uses final CAL calib., high forward tracking eff.
  - ◆ Output of production is immediately useful
- ◆ Major upgrade of Simulation
  - ◆ Detector geometry description, drift models, parameterized charge deposition models
  - ◆ Much wider use of “realistic” simulation
- ◆ Split and conquer
  - ◆ Run 2 datasets are split into 42 different physics streams (for example high pt muons = 3M events).
  - ◆ This is costly in farms operations but then reprocessing and physics analysis is much more efficient



Raw CDF Run2 Preliminary



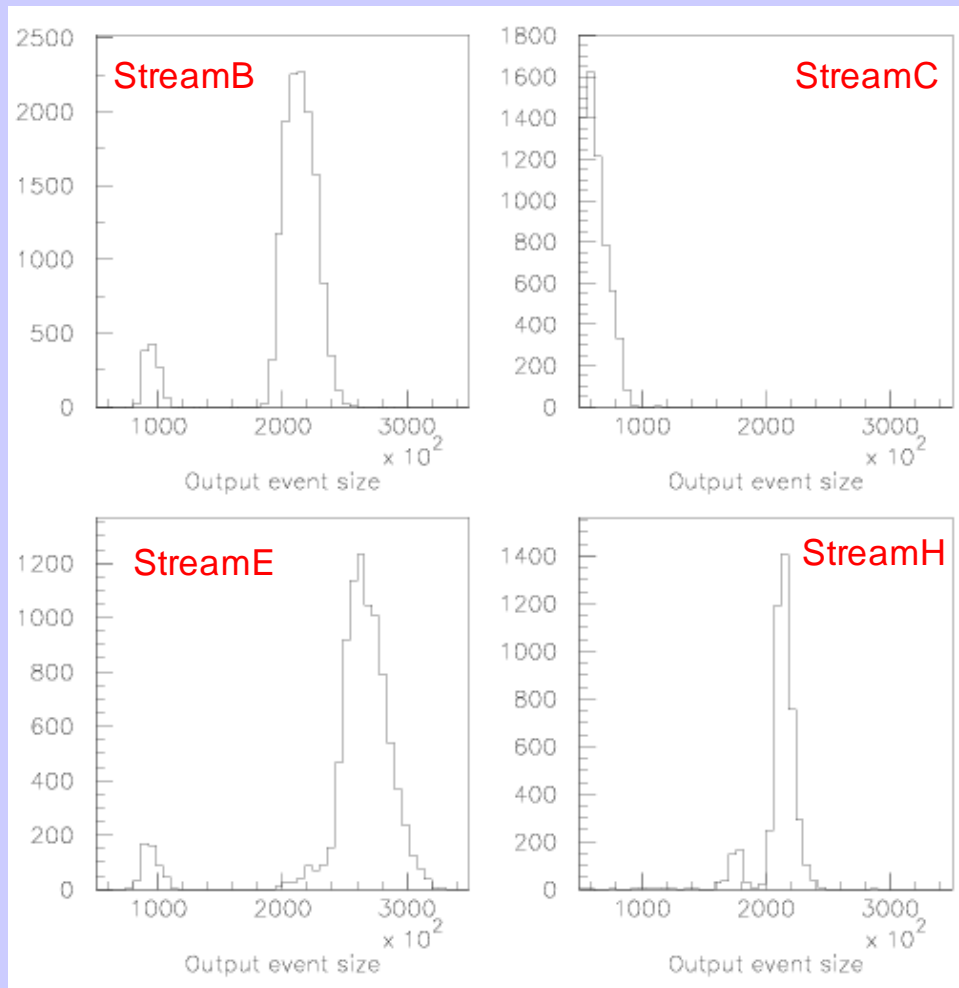


## ***What we plan to do***

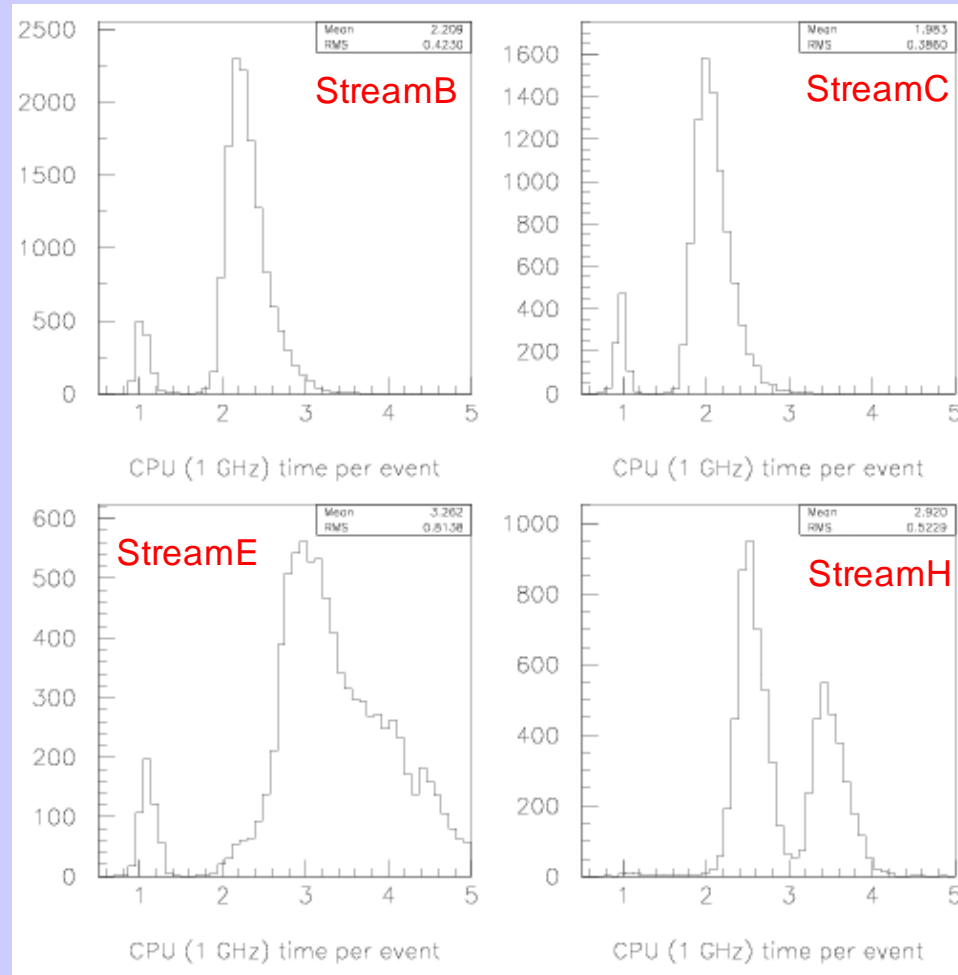
- ◆ Improve data format further
- ◆ Reduce the number of copies of data
- ◆ Continue to improve tracking
- ◆ Complete migration away from KAI compiler
- ◆ Prepare code base for required GRID infrastructure without disrupting physics
- ◆ Continue to support core software packages
- ◆ Improve support for analysis packages eg. Stntuple, Btag\*. They should be validated for frozen releases as they mature.

# ***Backup Slides***

## ***Output Event Size***

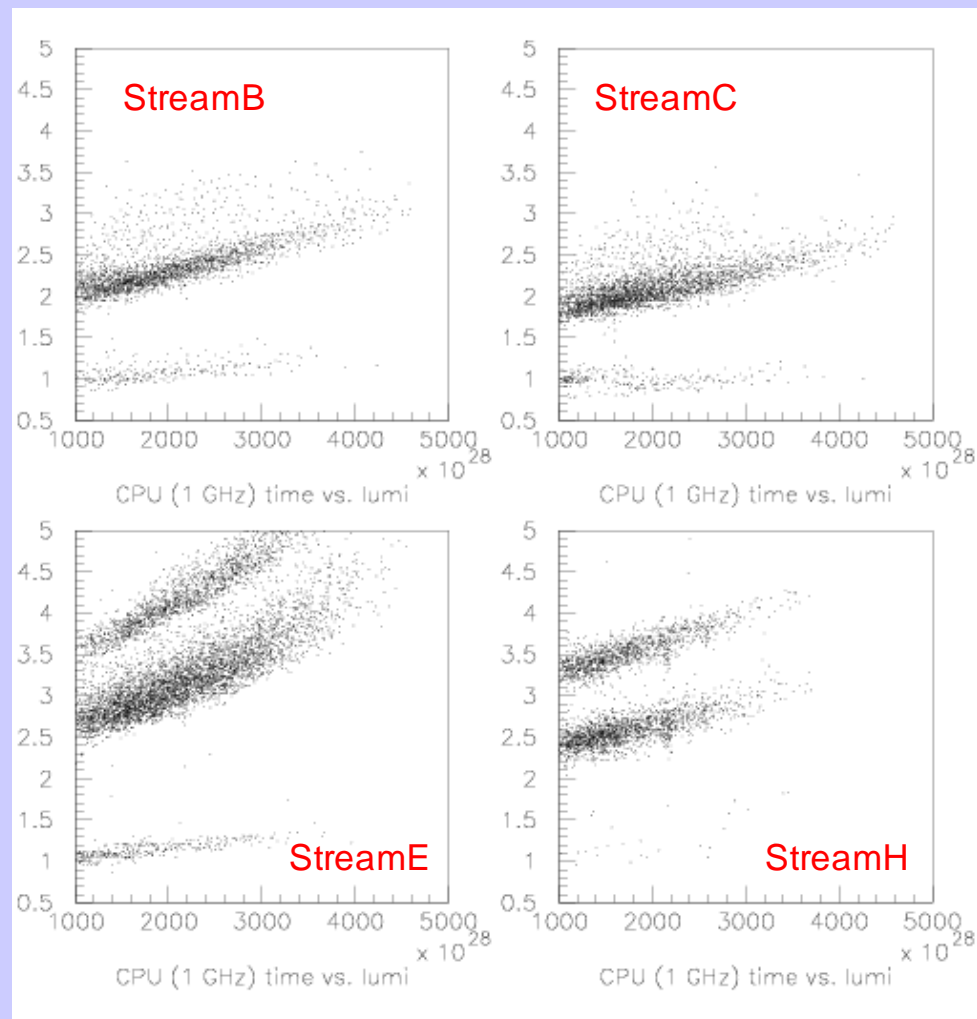


# *CPU time/event*





# *CPU vs. Luminosity*



# *CDF Detector*

